Implementation of Machine Learning Algorithm in Eye-Gaze Tracking using WSN based Convolutional Neural Networks

Jai Prakash Prasad

Associate Professor, Department of Electronics and Communication Engineering, Don Bosco Institute of Technology, Kumbalgodu, Mysore Road, Bangalore 560074, Karnataka, India

---

Abstract - *A paradigm shift is required to prevent the increasing automobile accident deaths that are mostly due to the inattentive behavior of drivers. Knowledge of gaze region can provide valuable information regarding a driver's point of attention. Accurate and inexpensive gaze classification systems in cars can improve safe driving. The margin of error in gaze detection increases when drivers gaze at objects by moving their eyes without moving their heads. Accurate and efficient eye gaze estimation is important for emerging consumer electronic systems such as driver monitoring systems and novel user interfaces. Such systems are required to operate reliably in difficult, unconstrained environments with low power consumption and at minimal cost. To solve this problem, a deep learning-based gaze estimation techniques, with a focus on WSN based Convolutional Neural Networks (CNN) based method has been considered. The proposed work provides architecture based on the data science perspective: First, a novel neural network architecture which is designed to exploit every possible visual attribute such as the states of both eyes and the head position, including several augmentations; secondly, the data fusion method is utilized by incorporating multiple gaze datasets. However, the accuracy of detecting and classification of the pupil center and corneal reflection center depends upon a car environment due to various conditions such as environment light changes, reflections on glasses surface, and motion and optical blurring of captured eye image. This work also implement pre-trained models, network architectures, and datasets that are useful for design and development of training CNN based deep learning models for Eye-Gaze Tracking and Classification.*

Key Words: Eye Gaze tracking, Convolutional Neural Network, Corneal Reflection, Deep Learning, Pupil Centre.

1. INTRODUCTION

Traffic accidents are mainly caused by a diminished driver vigilance level and gaze distraction from the road [1, 2]. Driver distraction is the main source of attention divergence from the roadway and can pose serious dangers to the lives of drivers, passengers, and pedestrians. According to the United States Department of Transportation, 3179 people were killed and 431,000 injured in 2017 due to distracted drivers [3]. Any activity that can divert driver attention from the primary task of driving can lead to distracted driving. It can happen for many reasons, but the most common are using a smart phone, controlling the radio, eating and drinking, and operating a global positing system (GPS). According to the National Highway Traffic Safety Administration (NHTSA) the risk factor for auto wrecks increases three times when drivers are using their smart phones during driving [4]. Using a smart phone causes the longest period of drivers taking their eyes off the road (EOR). In short, it can be a reason for driver distraction, and the technology of driver gaze detection can play a pivotal role in helping to avoid auto accidents. The classification of driver gaze attention is an area of increasing relevance in the pursuit of accident reduction.

Research for eye tracking techniques has been progressively implemented in many applications, such as driving fatigue-warning systems, mental health monitoring, eye-tracking controlled wheelchair, and other human–computer interface systems. However, there are several constraints such as reliable real-time performance, high accuracy, availability of components, and having a portable and non-intrusive system. It is also crucial to achieve higher system robustness against encountered challenges, such as changing light conditions, physical eye appearance, surrounding eye features, and reflections of eye-glasses.

Several related works have proposed eye-controlled wheelchair systems; however, these rarely address the constraints of the system's software performance, physical, and surrounding challenges beyond the system, novelty of algorithms, and user's comfort and safety altogether. Furthermore, the Convolutional Neural Network (CNNs) is a state-of-the-art and powerful tool that enables solving computationally and data-intensive problems. CNN is pioneering in a wide spectrum of applications in the context of object classification, speech recognition, and natural language processing and even wheelchair control, a more detailed literature review will be discussed in the later sections. However, the paper lacks high accuracy, real-time application, and does not provide the details of such a design, which could be useful for further improvement [12][13]. Although various studies concerned with gaze interaction have been reported, the performance of these methods in smart interactive environments under real-life conditions is still not satisfactory.

To solve this problem, a deep learning-based gaze estimation techniques, with a focus on Convolutional Neural Networks (CNN) based method has been considered. The proposed work provides architecture based on the data science perspective: First, a novel neural network architecture which is designed to exploit every possible visual attribute such as the states of both eyes and the head position, including several augmentations; secondly, the data fusion method is utilized by incorporating multiple gaze datasets. This work also implement pre-trained models, network architectures, and datasets that are useful for design and development of training CNN based deep learning models for Eye-Gaze Tracking and Classification.

2. LITERATURE SURVEY

Choi et al. [1] proposed a five-layered CNN-based technique for driver gaze zone classification and head-pose estimation. They created a dataset consisting of different images of male and female drivers, including drivers using eyeglasses. Based on the dataset, they built a CNN model that can classify 9 gaze zones of drivers and estimate their head-pose. Each gaze zone represents different regions in a car, including left mirror, right mirror, rearview mirror, steering, gear, middle, left windscreen, and right windscreen.

Naqvi, et al. [2] tackled this challenge by introducing a CNN-based model using a near-infrared (NIR) camera that considers head and eye movements and does not obstruct the view of drivers. The designed system is made up of one NIR camera, one zoom lens and six NIR light-emitting diodes (LEDs) for illumination. The NIR camera is used to capture the frontal view image of the driver, which is then transmitted to a laptop through a USB interface line.

Konrad, et al. [3] addressed this problem by introducing an end-to-end CNN-based gaze estimation technique for near-eye displays. They created a dataset containing eye images of users looking at various calibration points on a screen. The images were captured by a camera placed very close to the face of different subjects. Based on the dataset, they built a simple CNN model (using the LeNet architecture) that takes the images of users as input and estimate the gaze direction of the users, based on the x and y coordinates on the screen. The authors treated the gaze estimation problem as a multi-class classification problem, where each class is treated as a point on the screen. The technique was evaluated on the captured dataset, and it produced an angular error of 6.7 degrees, which is poor. The poor performance is likely due to the poor image quality (28×28) and the dataset variability used to train the network. The dataset contains images from only 5 subjects.

3. METHODOLOGY

Due to unrestricted user's head movement, head pose could be the additional feature to the training model. Hence, the orientation of head pose is extracted and represented by the Euler's angles: pitch, yaw, and roll as shown in figure 1.

Head Pose Estimation

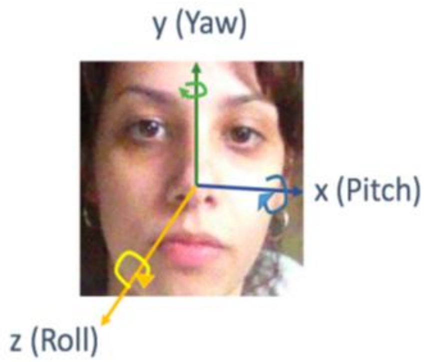


Figure 1: Head Pose using Euler's Angles; Pitch, yaw and roll

Training Details

Metrics: This metric is provided in the GazeCapture dataset.

Loss function: Euclidean's distance between predicted and target.

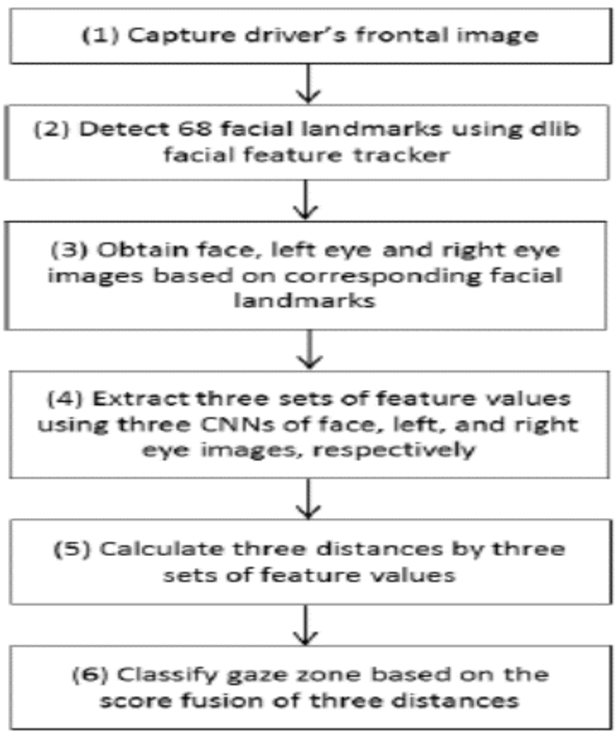
Validation: Some percent of data are used as validation dataset.

Learning rate: Initializing learning rate.

Optimizer: Adams optimizer.

Epochs: epochs count.

4. FLOWCHART



5. CONCLUSIONS

We are proposing a method of driver gaze classification in the vehicular environment based on CNN. For driver gaze classification, face, left eye, and right eye images will be obtained from input image based on the ROI (region-of-interest) defined by facial landmarks from the facial feature tracker. Further fine tuning with a pre-trained CNN model separately for the extracted cropped images of face, left eye, and right eye using VGG-face network to obtain the required gaze features from the fully connected layer of the network. Three distances based on all the obtained features are combined to find the final result of classification. The impact of CNN model on gaze classification will be also studied.

FUTURE ENHANCEMENT

Our method can also be applied to the images by visible light camera without an additional illuminator, which was proved by the experiments with open Columbia gaze dataset CAVE-DB.

APPLICATION

Research for eye tracking techniques has been progressively implemented in many applications, such as driving fatigue-warning systems, mental health monitoring, eye-tracking controlled wheelchair, and other human–computer interface systems.

REFERENCES

[1] H. Choi, Y. G. Kim, and T. B. H. Tran, ''Real-time categorization of driver's gaze zone and head pose - using the convolutional neural network,'' in Proc. HCI Korea, Jan. 2016, pp. 417–422.

[2] R. Naqvi, M. Arsalan, G. Batchuluun, H. Yoon, and K. Park, ''Deep learning-based gaze detection system for automobile drivers using a NIR camera sensor,'' Sensors, vol. 18, no. 2, p. 456, Feb. 2018.

[3] R. Konrad, S. Shrestha, and P. Varma, ''Near-eye display gaze tracking via convolutional neural networks,'' Standford Univ., Standford, CA, USA, Tech. Rep., 2016.

[4] S. Hickson, N. Dufour, A. Sud, V. Kwatra, and I. Essa, ''Eyemotion: Classifying facial expressions in VR using eye-tracking cameras,'' in Proc. IEEE Winter Conf. Appl. Comput. Vis. (WACV), Jan. 2019, pp. 1626–1635.

[5] S. J. Garbin, O. Komogortsev, R. Cavin, G. Hughes, Y. Shen, I. Schuetz, and S. S. Talathi, ''Dataset for eye tracking on a virtual reality platform,'' in Proc. Symp. Eye Tracking Res. Appl., Jun. 2020, pp. 1–10.

[6] Jai Prakash Prasad and Dr. S C Mohan, "Security Enhancement of N-Tier Grid Protocol Using Elliptical Curve Cryptography for WSN", IRACST – International Journal of Computer Networks and Wireless Communications (IJCNWC), Volume 5, No. 6, Page No. 651-656, ISSN: 2250 - 3501 (Online), 2277 - 5307 (Print), Dec 2015, Impact Factor: 0.765.

[7] Jai Prakash Prasad and Dr. S C Mohan, "Dynamic GRID Clustered Protocol to Improvise QOS Performance of Wireless Sensor Networks", International Journal of Computing, Communications and Networking (IJCCN), Volume 4, No.4, Page No. 13-17, ISSN Online: 2319 – 2720, Oct-Nov 2015, Impact Factor: 1.13.

[8] Jai Prakash Prasad and Dr. S C Mohan, "Spherical Grid Protocol to Enhance Quality of Service to Resource Constrained Wireless Sensor Networks", International Journal of Sensors and Sensor Networks (IJSSN), Volume 4, Issue Number 1, Page No. 1-6, DOI:10.11648/j.ijssn.20160401.11, 14 Feb 2016.

[9] Jai Prakash Prasad and Dr. S C Mohan, "Energy Efficient Elliptical Curve based Spherical Grid Routing Protocol for Wireless Sensor Networks", Global Journals of Computer Science and Technology (GJCST), Volume 16, Issue 3, Page No. 18-25, Version 1.0 Year 2016. Page No. 19-25, Online ISSN: 0975-4172 & Print ISSN: 0975-4350.

[10] Jai Prakash Prasad and Dr. S C Mohan, "A Novel Energy Efficient Spherical Routing Model for Dynamic Route Discovery in Wireless Sensor Networks", Published in International Journal of Engineering Research & Technology (IJERT), Vol. 4, Issue No. 29, Page No. 1-6, ISSN: 2278-0181, Year 2016.

[11] Jai Prakash Prasad and Dr. S C Mohan, "Optimized Implementation of Energy Efficient Spherical Grid Routing to Monitor Fast Moving Object", GRENZE International Journal of Engineering and Technology (GIJET), Grenze ID: 01.GIJET.3.3.10, Pages: 15-19, Volume 3, Published Year: 2017.